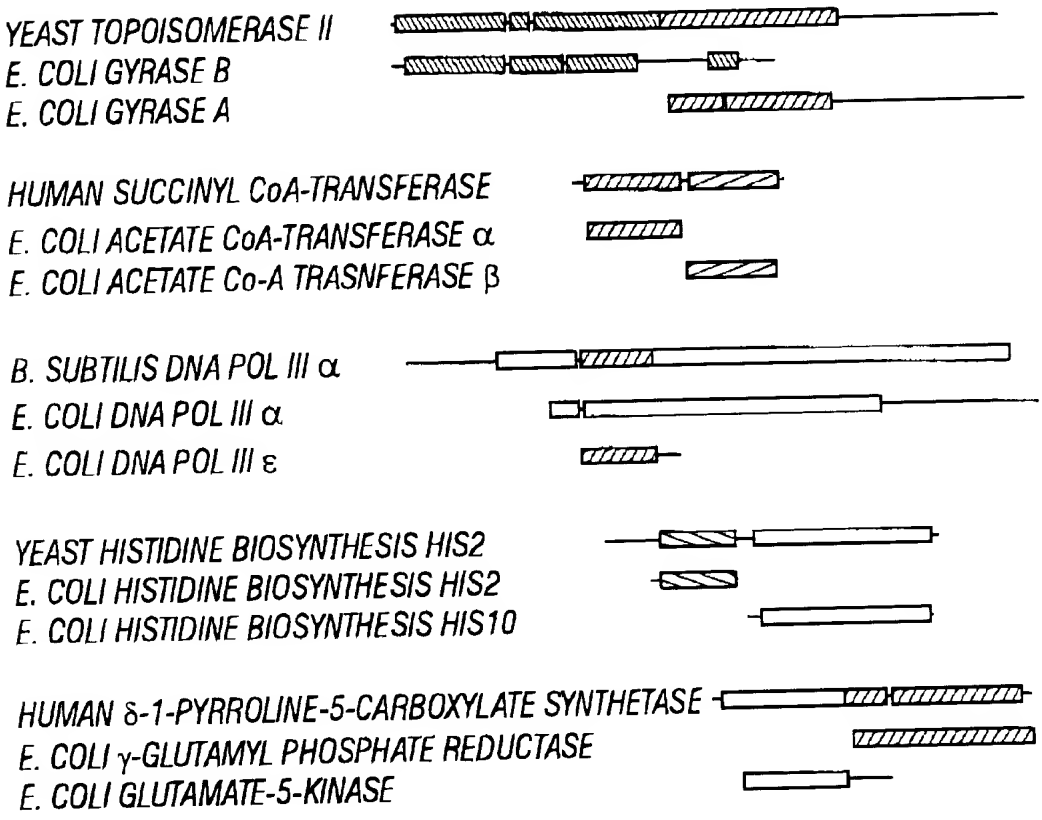
YEAST TOPOISOMERASE II
E. COLI GYRASE B
E. COLI GYRASE A
HUMAN SUCCINYL CoA-TRANSFERASE
E. COLI ACETATE CoA-TRANSFERASE α
E. COLI ACETATE Co-A TRASNFERASE β
B. SUBTILIS DNA POL III α
E. COLI DNA POL III α
E. COLI DNA POL III ε
YEAST HISTIDINE BIOSYNTHESIS HIS2
E. COLI HISTIDINE BIOSYNTHESIS HIS2
E. COLI HISTIDINE BIOSYNTHESIS HIS10
HUMAN δ-1-PYRROLINE-5-CARBOXYLATE SYNTHETASE
E. COLI γ-GLUTAMYL PHOSPHATE REDUCTASE
E. COLI GLUTAMATE-5-KINASE

FIG. 1A

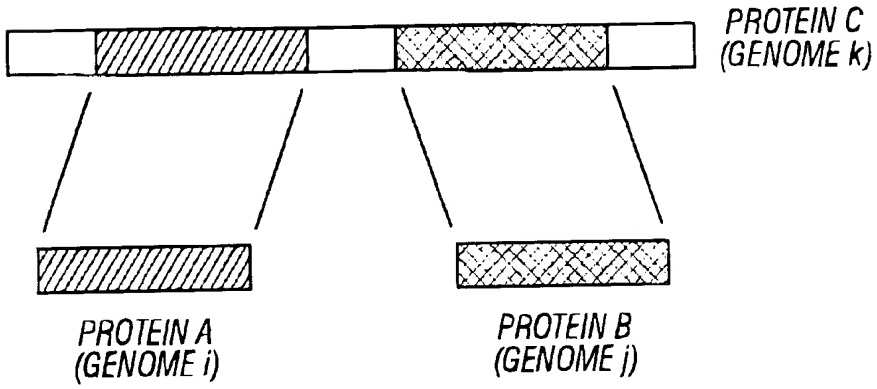
PROTEIN C
(GENOME k)
PROTEIN A
(GENOME i)
PROTEIN B
(GENOME j)

FIG. 1B

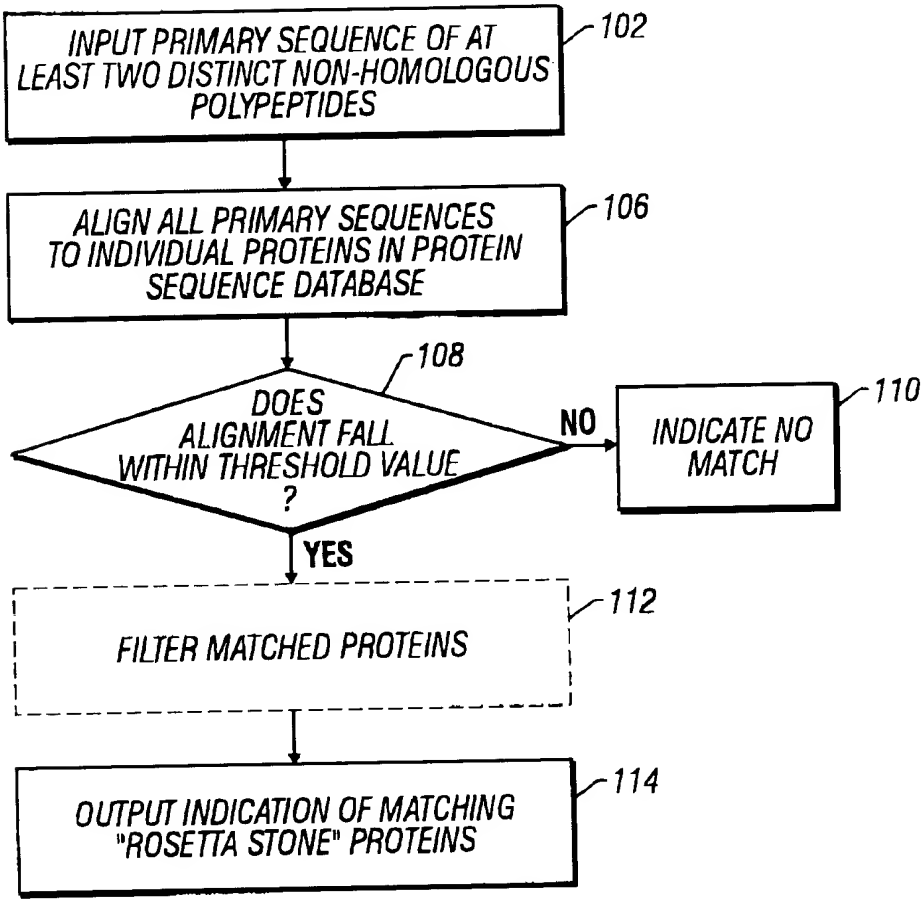
INPUT PRIMARY SEQUENCE OF AT LEAST TWO DISTINCT NON-HOMOLOGOUS POLYPEPTIDES
102
ALIGN ALL PRIMARY SEQUENCES TO INDIVIDUAL PROTEINS IN PROTEIN SEQUENCE DATABASE
106
108
DOES ALIGNMENT FALL WITHIN THRESHOLD VALUE ?
NO
INDICATE NO MATCH
110
YES
FILTER MATCHED PROTEINS
112
OUTPUT INDICATION OF MATCHING "ROSETTA STONE" PROTEINS
114

FIG. 2A

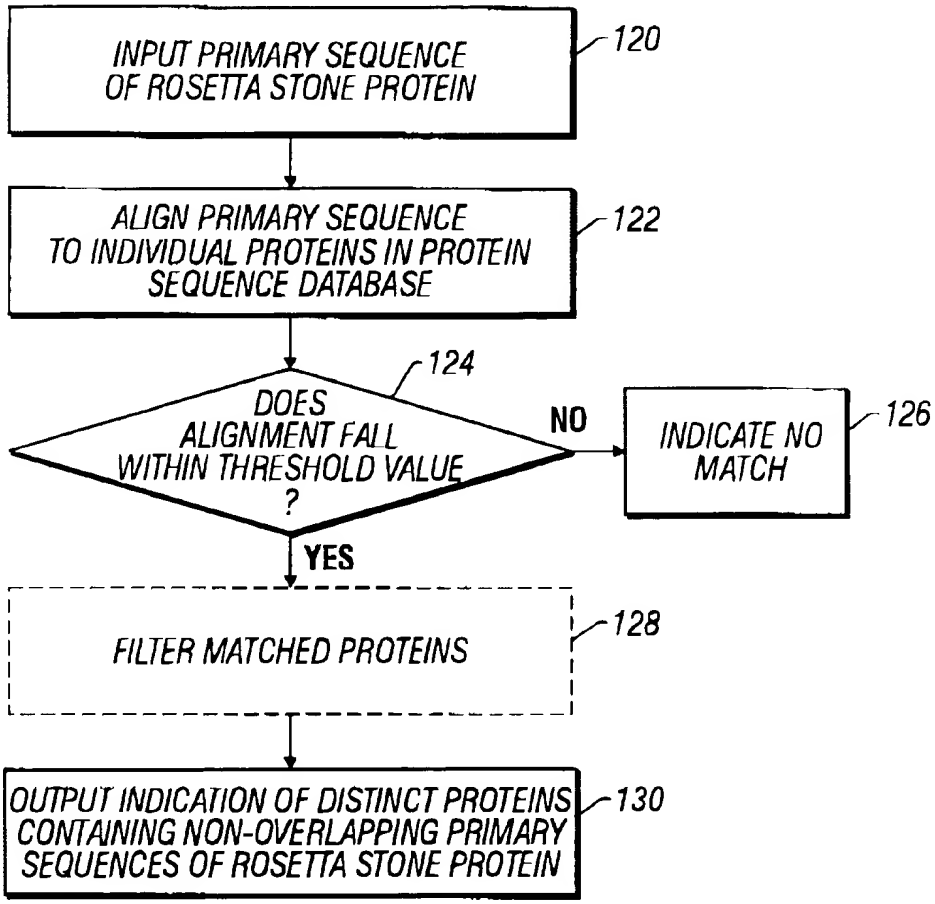
INPUT PRIMARY SEQUENCE OF ROSETTA STONE PROTEIN
120
ALIGN PRIMARY SEQUENCE TO INDIVIDUAL PROTEINS IN PROTEIN SEQUENCE DATABASE
122
124
DOES ALIGNMENT FALL WITHIN THRESHOLD VALUE ?
NO
INDICATE NO MATCH
126
YES
FILTER MATCHED PROTEINS
128
OUTPUT INDICATION OF DISTINCT PROTEINS CONTAINING NON-OVERLAPPING PRIMARY SEQUENCES OF ROSETTA STONE PROTEIN
130

FIG. 2B

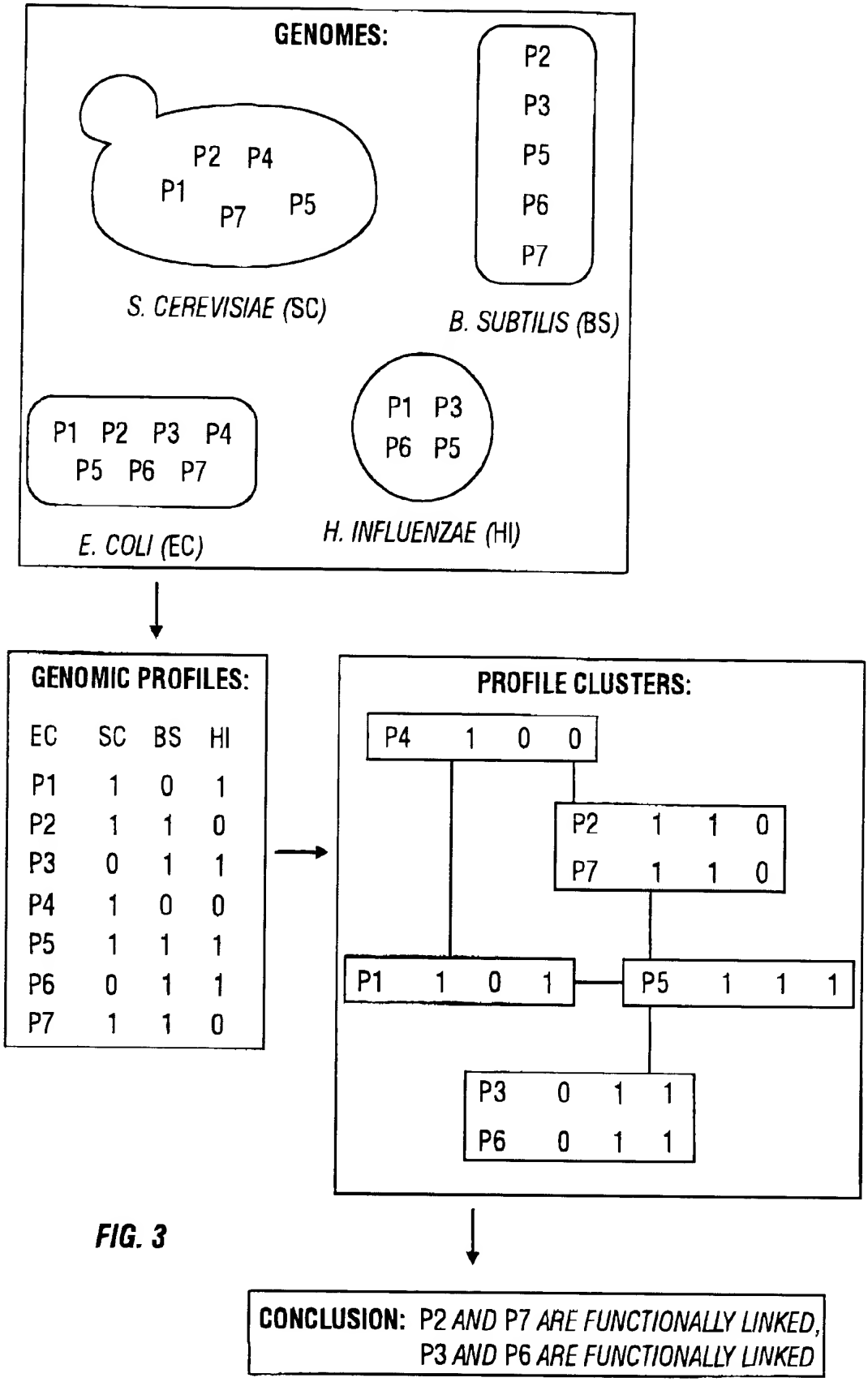
GENOMES:
P2 P4
P1
P7 P5
S. CEREVISIAE (SC)
P2
P3
P5
P6
P7
B. SUBTILIS (BS)
P1 P2 P3 P4
P5 P6 P7
E. COLI (EC)
P1 P3
P6 P5
H. INFLUENZAE (HI)
GENOMIC PROFILES:
EC SC BS HI
P1 1 0 1
P2 1 1 0
P3 0 1 1
P4 1 0 0
P5 1 1 1
P6 0 1 1
P7 1 1 0
PROFILE CLUSTERS:
P4 1 0 0
P2 1 1 0
P7 1 1 0
P1 1 0 1
P5 1 1 1
P3 0 1 1
P6 0 1 1
CONCLUSION: P2 AND P7 ARE FUNCTIONALLY LINKED,
P3 AND P6 ARE FUNCTIONALLY LINKED

FIG. 3

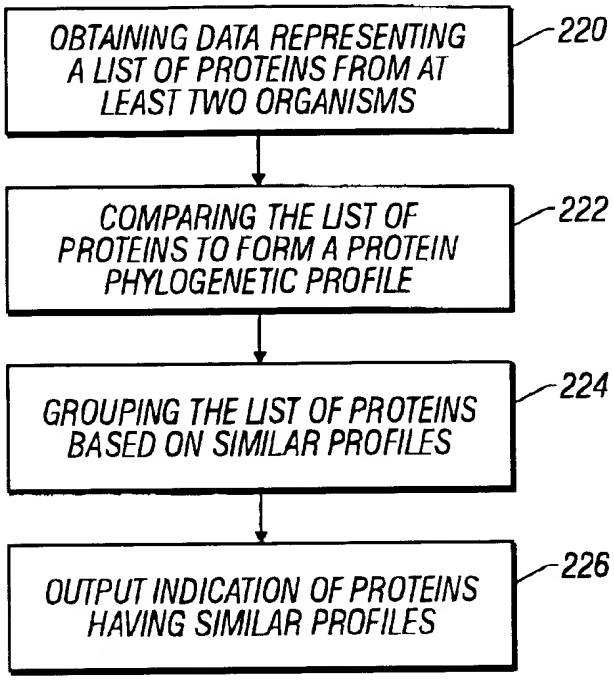
OBTAINING DATA REPRESENTING A LIST OF PROTEINS FROM AT LEAST TWO ORGANISMS
220
COMPARING THE LIST OF PROTEINS TO FORM A PROTEIN PHYLOGENETIC PROFILE
222
GROUPING THE LIST OF PROTEINS BASED ON SIMILAR PROFILES
224
OUTPUT INDICATION OF PROTEINS HAVING SIMILAR PROFILES
226

FIG. 4A

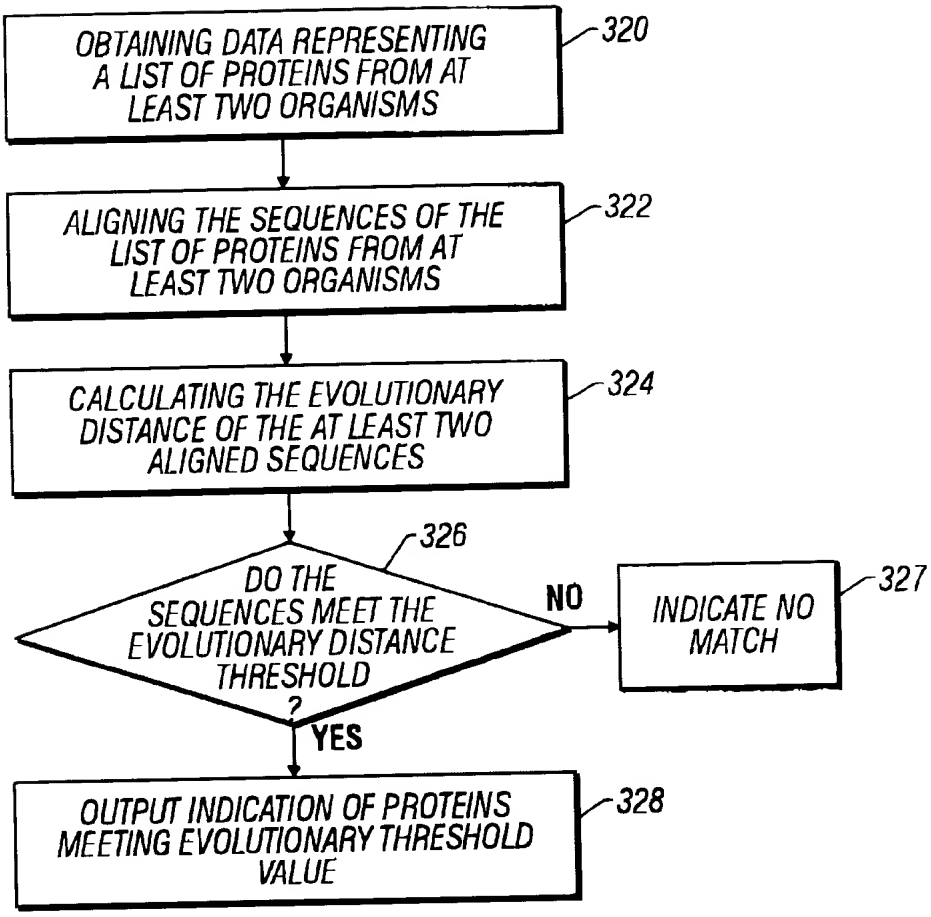
OBTAINING DATA REPRESENTING A LIST OF PROTEINS FROM AT LEAST TWO ORGANISMS
320
ALIGNING THE SEQUENCES OF THE LIST OF PROTEINS FROM AT LEAST TWO ORGANISMS
322
CALCULATING THE EVOLUTIONARY DISTANCE OF THE AT LEAST TWO ALIGNED SEQUENCES
324
326
DO THE SEQUENCES MEET THE EVOLUTIONARY DISTANCE THRESHOLD ?
NO
INDICATE NO MATCH
327
YES
OUTPUT INDICATION OF PROTEINS MEETING EVOLUTIONARY THRESHOLD VALUE
328

FIG. 4B

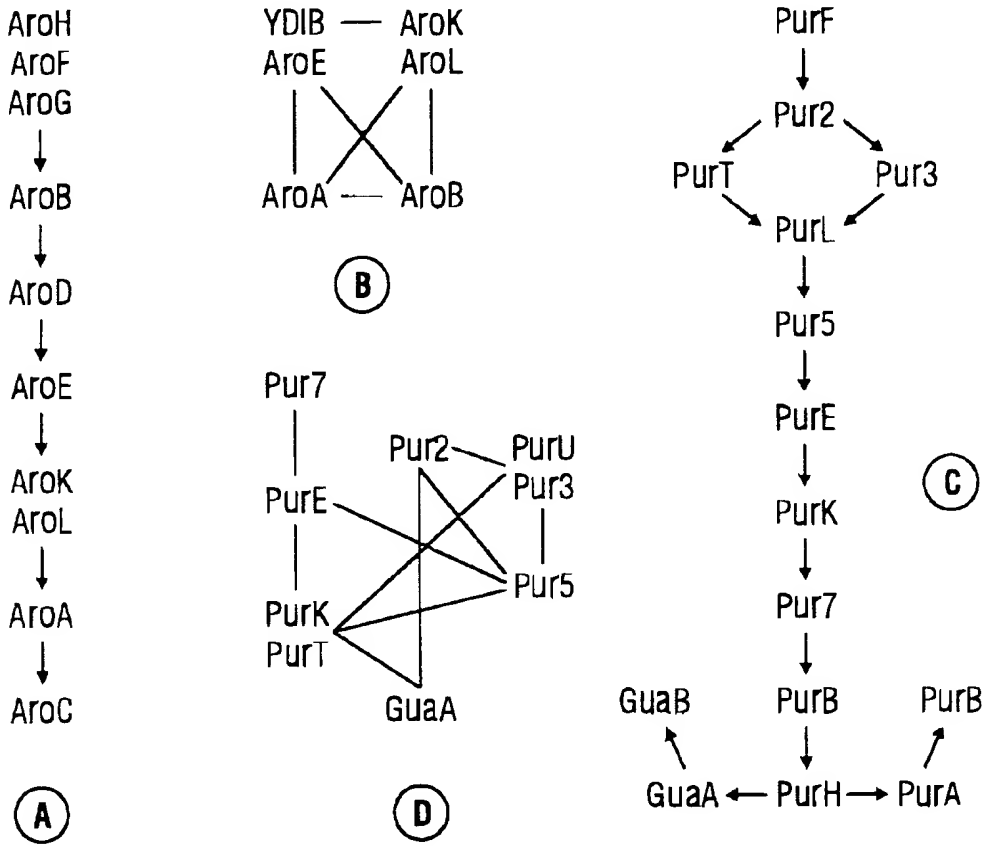
AroH
AroF
AroG
AroB
AroD
AroE
AroK
AroL
AroA
AroC
A
YDIB
AroK
AroE
AroL
AroA
AroB
B
Pur7
Pur2
PurU
PurE
Pur3
Pur5
PurK
PurT
GuaA
D
PurF
Pur2
PurT
Pur3
PurL
Pur5
PurE
PurK
Pur7
PurB
C
GuaB
PurB
PurB
GuaA
PurH
PurA

FIG. 5

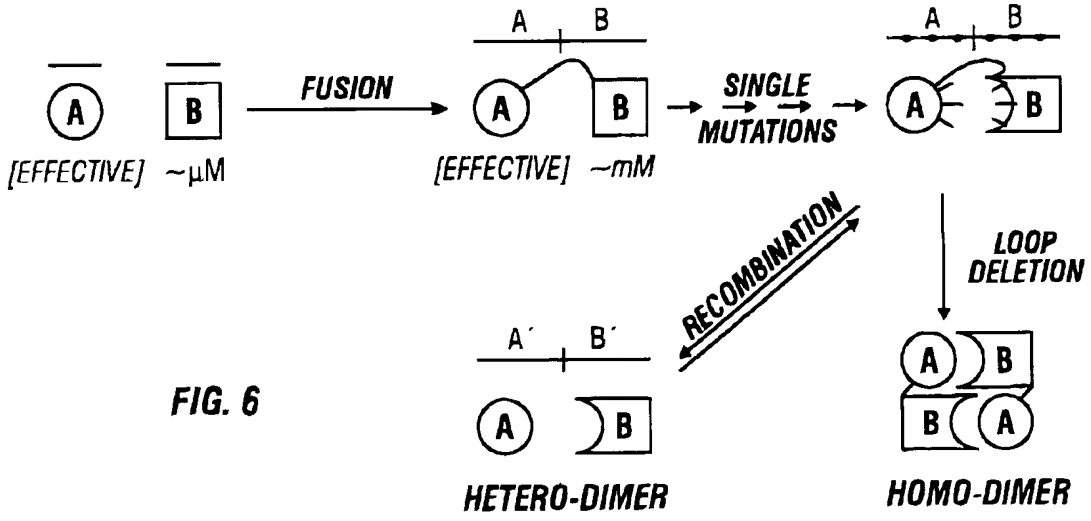
A
B
[EFFECTIVE]
~μM
FUSION
A
B
[EFFECTIVE]
~mM
SINGLE
MUTATIONS
A
B
LOOP
DELETION
RECOMBINATION
A´
B´
A
B
HETERO-DIMER
A
B
B
A
HOMO-DIMER

FIG. 6

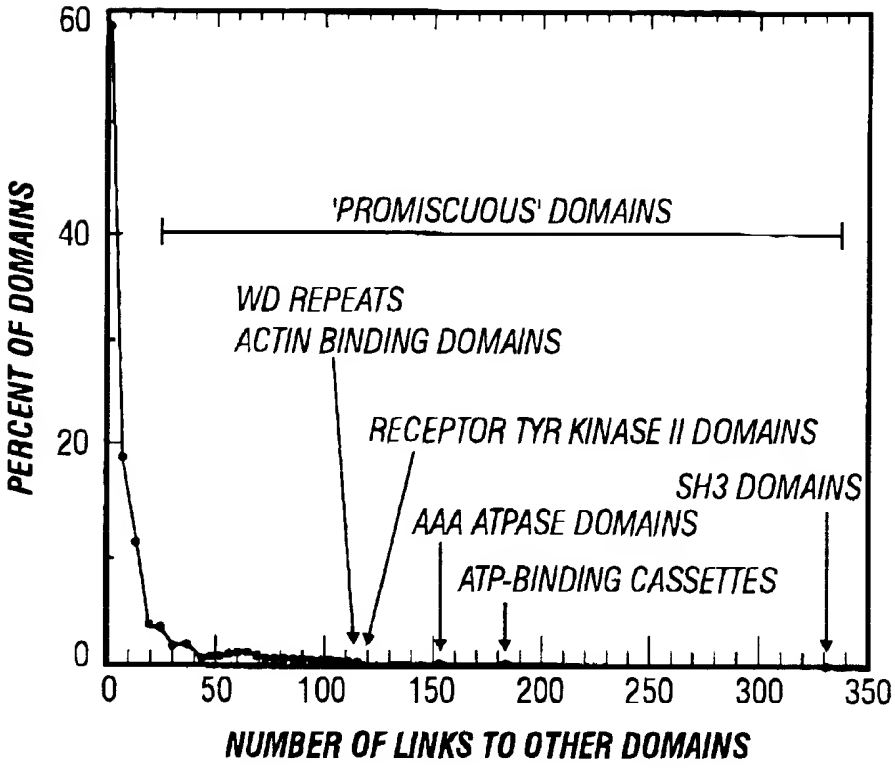
60
40
20
0
PERCENT OF DOMAINS
'PROMISCUOUS' DOMAINS
WD REPEATS
ACTIN BINDING DOMAINS
RECEPTOR TYR KINASE II DOMAINS
SH3 DOMAINS
AAA ATPASE DOMAINS
ATP-BINDING CASSETTES
0
50
100
150
200
250
300
350
NUMBER OF LINKS TO OTHER DOMAINS

FIG. 7

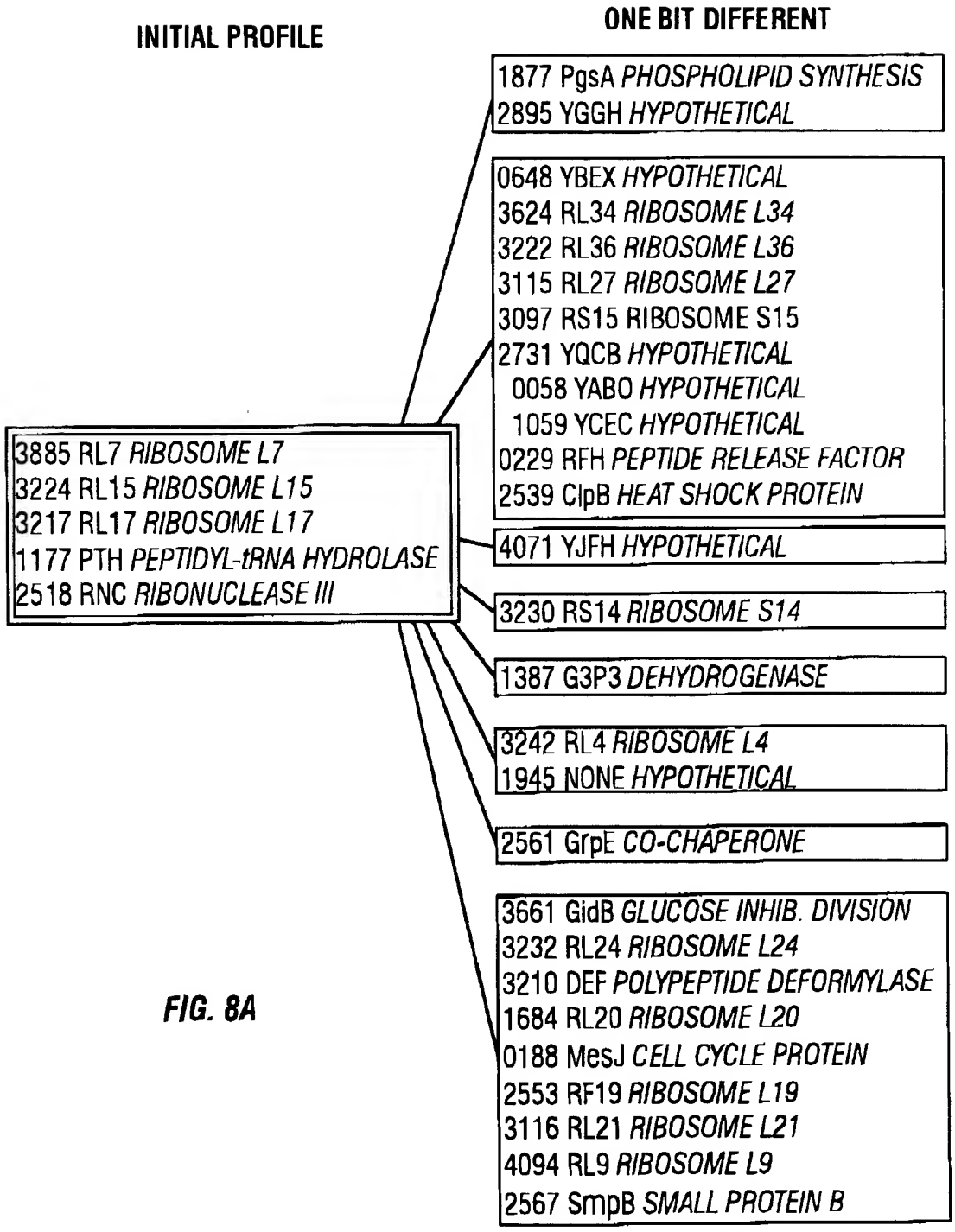
INITIAL PROFILE
ONE BIT DIFFERENT
3885 RL7 RIBOSOME L7
3224 RL15 RIBOSOME L15
3217 RL17 RIBOSOME L17
1177 PTH PEPTIDYL-tRNA HYDROLASE
2518 RNC RIBONUCLEASE III
1877 PgsA PHOSPHOLIPID SYNTHESIS
2895 YGGH HYPOTHETICAL
0648 YBEX HYPOTHETICAL
3624 RL34 RIBOSOME L34
3222 RL36 RIBOSOME L36
3115 RL27 RIBOSOME L27
3097 RS15 RIBOSOME S15
2731 YQCB HYPOTHETICAL
0058 YABO HYPOTHETICAL
1059 YCEC HYPOTHETICAL
0229 RFH PEPTIDE RELEASE FACTOR
2539 ClpB HEAT SHOCK PROTEIN
4071 YJFH HYPOTHETICAL
3230 RS14 RIBOSOME S14
1387 G3P3 DEHYDROGENASE
3242 RL4 RIBOSOME L4
1945 NONE HYPOTHETICAL
2561 GrpE CO-CHAPERONE
3661 GidB GLUCOSE INHIB. DIVISION
3232 RL24 RIBOSOME L24
3210 DEF POLYPEPTIDE DEFORMYLASE
1684 RL20 RIBOSOME L20
0188 MesJ CELL CYCLE PROTEIN
2553 RF19 RIBOSOME L19
3116 RL21 RIBOSOME L21
4094 RL9 RIBOSOME L9
2567 SmpB SMALL PROTEIN B

FIG. 8A

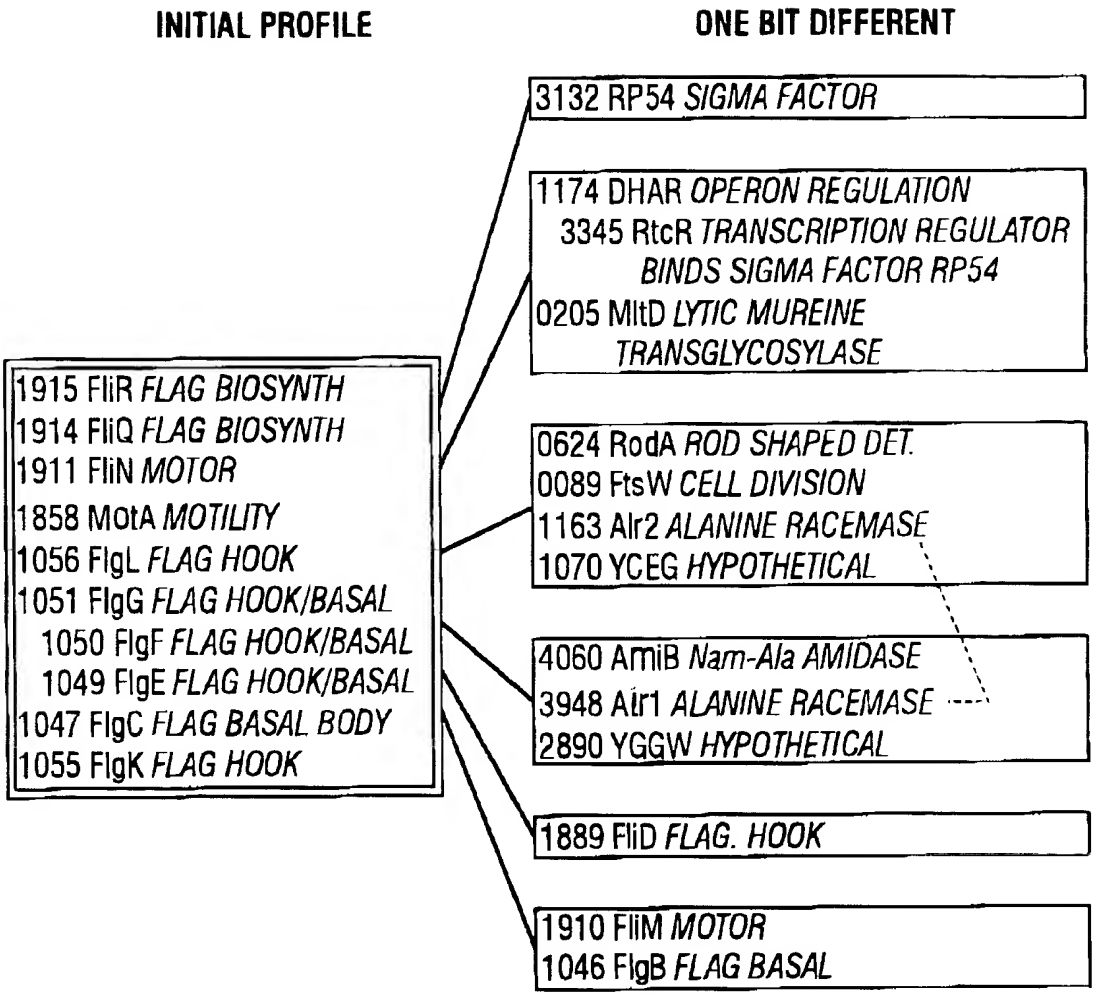
INITIAL PROFILE
ONE BIT DIFFERENT
1915 FliR FLAG BIOSYNTH
1914 FliQ FLAG BIOSYNTH
1911 FliN MOTOR
1858 MotA MOTILITY
1056 FlgL FLAG HOOK
1051 FlgG FLAG HOOK/BASAL
1050 FlgF FLAG HOOK/BASAL
1049 FlgE FLAG HOOK/BASAL
1047 FlgC FLAG BASAL BODY
1055 FlgK FLAG HOOK
3132 RP54 SIGMA FACTOR
1174 DHAR OPERON REGULATION
3345 RtcR TRANSCRIPTION REGULATOR BINDS SIGMA FACTOR RP54
0205 MltD LYTIC MUREINE TRANSGLYCOSYLASE
0624 RodA ROD SHAPED DET.
0089 FtsW CELL DIVISION
1163 Alr2 ALANINE RACEMASE
1070 YCEG HYPOTHETICAL
4060 AmiB Nam-Ala AMIDASE
3948 Alr1 ALANINE RACEMASE
2890 YGGW HYPOTHETICAL
1889 FliD FLAG. HOOK
1910 FliM MOTOR
1046 FlgB FLAG BASAL

FIG. 8B

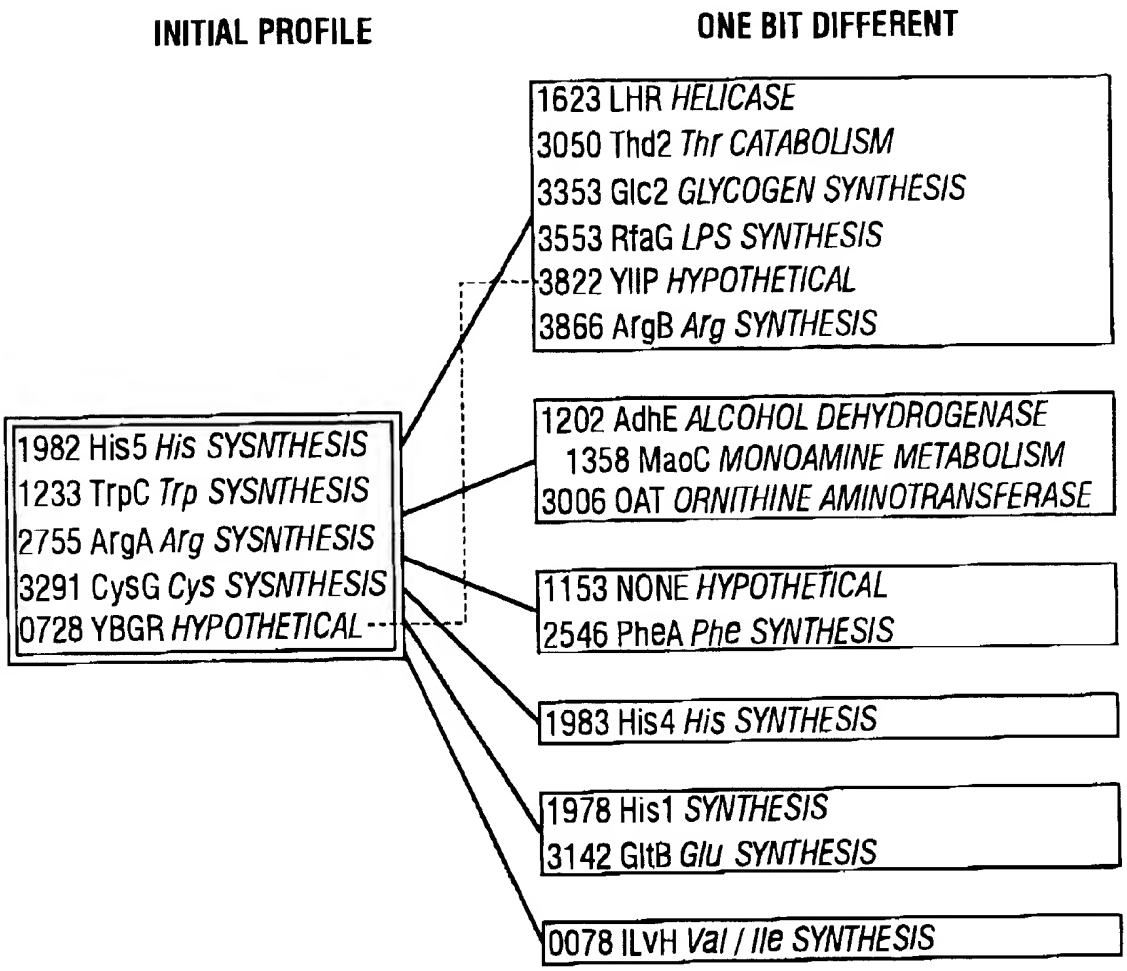
INITIAL PROFILE
ONE BIT DIFFERENT
1982 His5 His SYSNTHESIS
1233 TrpC Trp SYSNTHESIS
2755 ArgA Arg SYSNTHESIS
3291 CysG Cys SYSNTHESIS
0728 YBGR HYPOTHETICAL
1623 LHR HELICASE
3050 Thd2 Thr CATABOLISM
3353 Glc2 GLYCOGEN SYNTHESIS
3553 RfaG LPS SYNTHESIS
3822 YIIP HYPOTHETICAL
3866 ArgB Arg SYNTHESIS
1202 AdhE ALCOHOL DEHYDROGENASE
1358 MaoC MONOAMINE METABOLISM
3006 OAT ORNITHINE AMINOTRANSFERASE
1153 NONE HYPOTHETICAL
2546 PheA Phe SYNTHESIS
1983 His4 His SYNTHESIS
1978 His1 SYNTHESIS
3142 GltB Glu SYNTHESIS
0078 ILVH Val / Ile SYNTHESIS

FIG. 8C

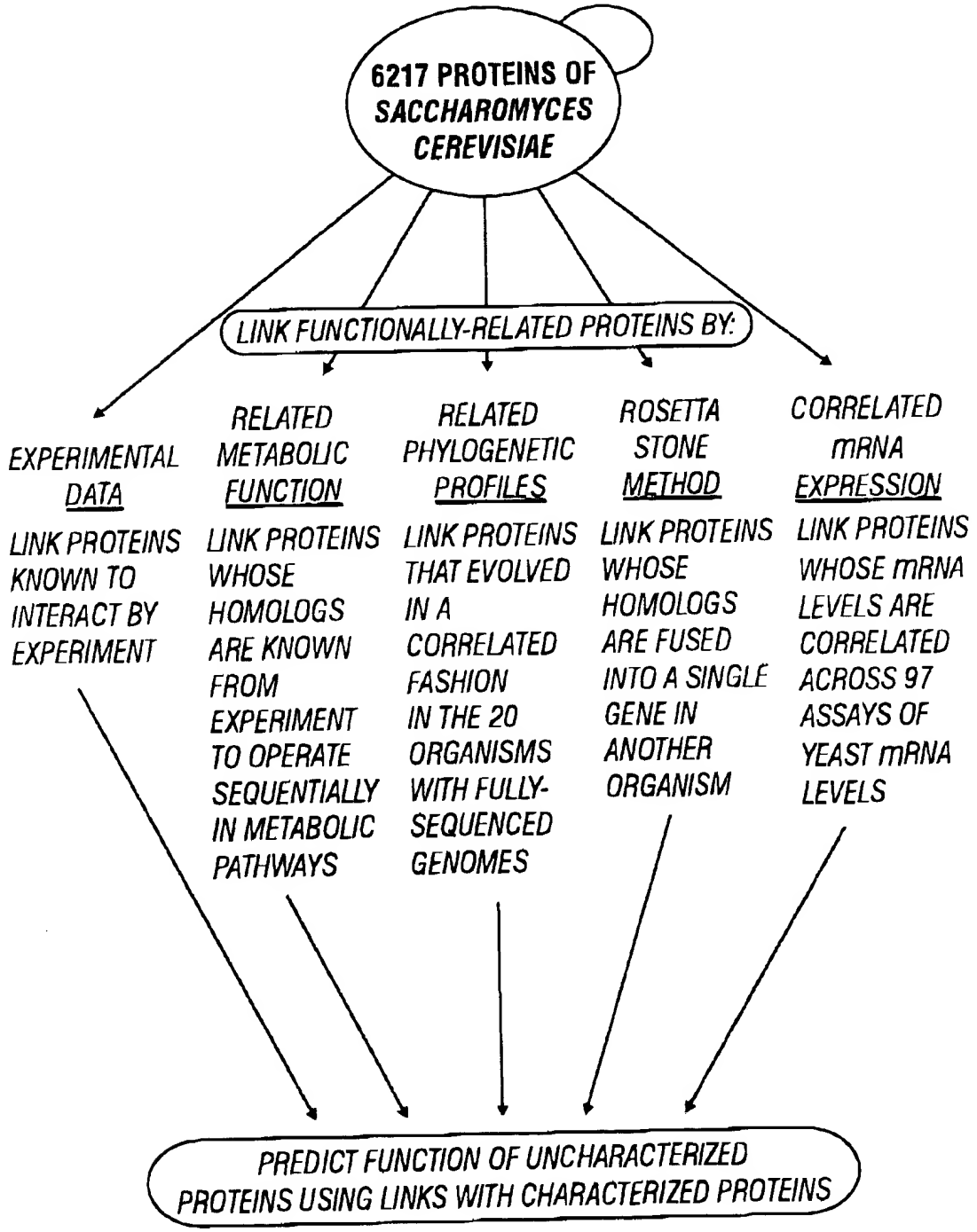
6217 PROTEINS OF SACCHAROMYCES CEREVISIAE
LINK FUNCTIONALLY-RELATED PROTEINS BY:
EXPERIMENTAL DATA
LINK PROTEINS KNOWN TO INTERACT BY EXPERIMENT
RELATED METABOLIC FUNCTION
LINK PROTEINS WHOSE HOMOLOGS ARE KNOWN FROM EXPERIMENT TO OPERATE SEQUENTIALLY IN METABOLIC PATHWAYS
RELATED PHYLOGENETIC PROFILES
LINK PROTEINS THAT EVOLVED IN A CORRELATED FASHION IN THE 20 ORGANISMS WITH FULLY-SEQUENCED GENOMES
ROSETTA STONE METHOD
LINK PROTEINS WHOSE HOMOLOGS ARE FUSED INTO A SINGLE GENE IN ANOTHER ORGANISM
CORRELATED mRNA EXPRESSION
LINK PROTEINS WHOSE mRNA LEVELS ARE CORRELATED ACROSS 97 ASSAYS OF YEAST mRNA LEVELS
PREDICT FUNCTION OF UNCHARACTERIZED PROTEINS USING LINKS WITH CHARACTERIZED PROTEINS

FIG. 9

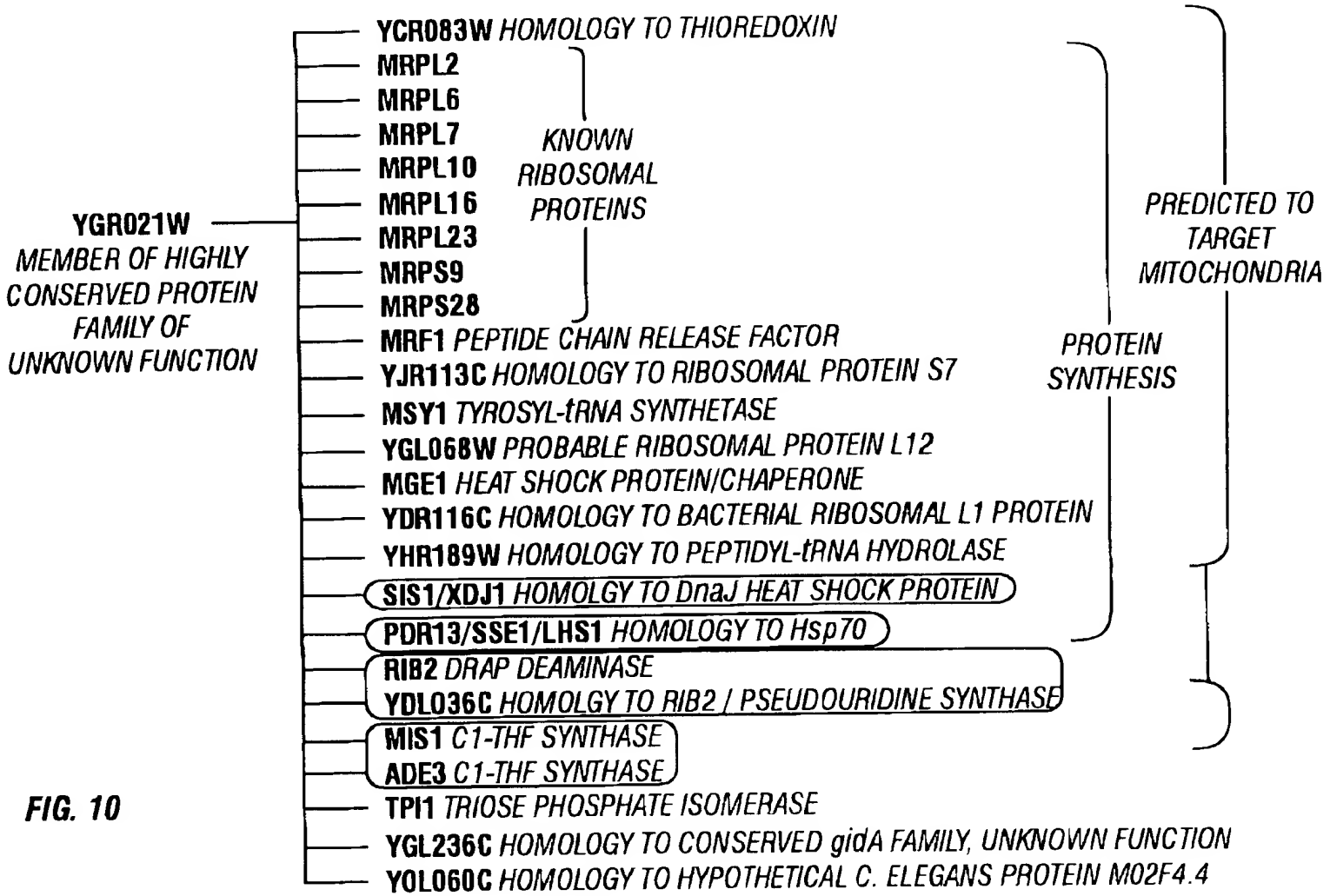
YGR021W
MEMBER OF HIGHLY CONSERVED PROTEIN FAMILY OF UNKNOWN FUNCTION
YCR083W HOMOLOGY TO THIOREDOXIN
MRPL2
MRPL6
MRPL7
MRPL10
MRPL16
MRPL23
MRPS9
MRPS28
KNOWN RIBOSOMAL PROTEINS
MRF1 PEPTIDE CHAIN RELEASE FACTOR
YJR113C HOMOLOGY TO RIBOSOMAL PROTEIN S7
MSY1 TYROSYL-tRNA SYNTHETASE
YGL068W PROBABLE RIBOSOMAL PROTEIN L12
MGE1 HEAT SHOCK PROTEIN/CHAPERONE
YDR116C HOMOLOGY TO BACTERIAL RIBOSOMAL L1 PROTEIN
YHR189W HOMOLOGY TO PEPTIDYL-tRNA HYDROLASE
SIS1/XDJ1 HOMOLGY TO DnaJ HEAT SHOCK PROTEIN
PDR13/SSE1/LHS1 HOMOLOGY TO Hsp70
RIB2 DRAP DEAMINASE
YDL036C HOMOLGY TO RIB2 / PSEUDOURIDINE SYNTHASE
MIS1 C1-THF SYNTHASE
ADE3 C1-THF SYNTHASE
TPI1 TRIOSE PHOSPHATE ISOMERASE
YGL236C HOMOLOGY TO CONSERVED gidA FAMILY, UNKNOWN FUNCTION
YOL060C HOMOLOGY TO HYPOTHETICAL C. ELEGANS PROTEIN M02F4.4
PROTEIN SYNTHESIS
PREDICTED TO TARGET MITOCHONDRIA

FIG. 10

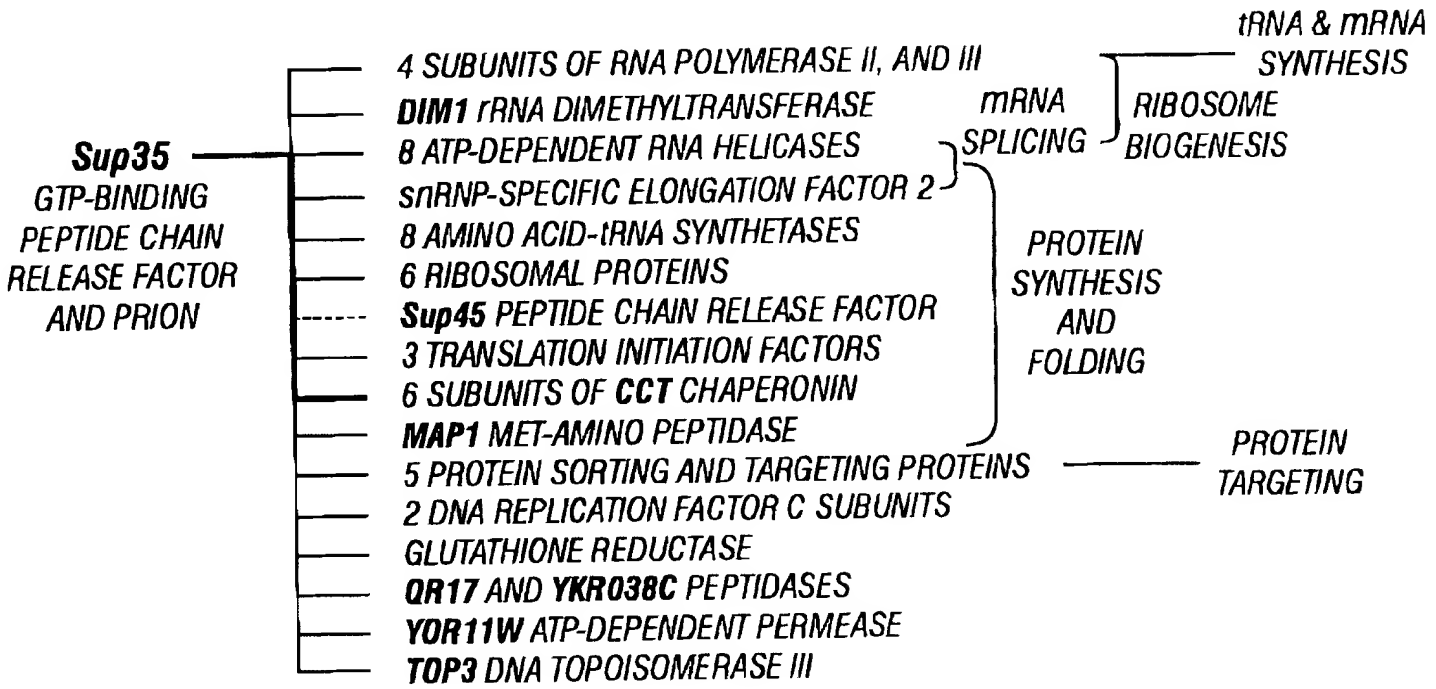
Sup35
GTP-BINDING
PEPTIDE CHAIN
RELEASE FACTOR
AND PRION
4 SUBUNITS OF RNA POLYMERASE II, AND III
DIM1 rRNA DIMETHYLTRANSFERASE
8 ATP-DEPENDENT RNA HELICASES
snRNP-SPECIFIC ELONGATION FACTOR 2
8 AMINO ACID-tRNA SYNTHETASES
6 RIBOSOMAL PROTEINS
Sup45 PEPTIDE CHAIN RELEASE FACTOR
3 TRANSLATION INITIATION FACTORS
6 SUBUNITS OF CCT CHAPERONIN
MAP1 MET-AMINO PEPTIDASE
5 PROTEIN SORTING AND TARGETING PROTEINS
2 DNA REPLICATION FACTOR C SUBUNITS
GLUTATHIONE REDUCTASE
QR17 AND YKR038C PEPTIDASES
YOR11W ATP-DEPENDENT PERMEASE
TOP3 DNA TOPOISOMERASE III
tRNA & mRNA
SYNTHESIS
mRNA
SPLICING
RIBOSOME
BIOGENESIS
PROTEIN
SYNTHESIS
AND
FOLDING
PROTEIN
TARGETING

FIG. 11A

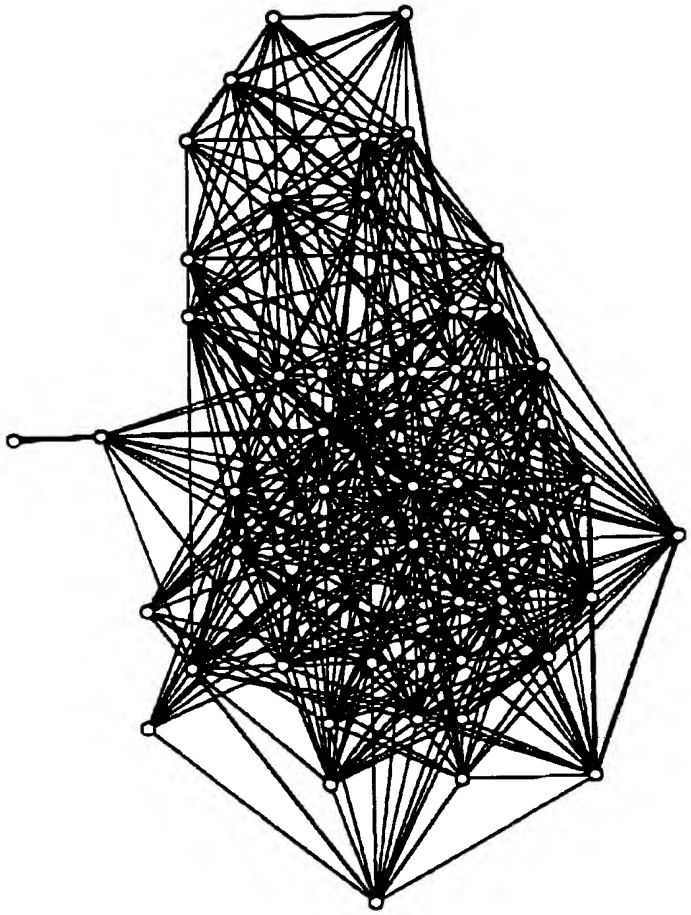

FIG. 11B

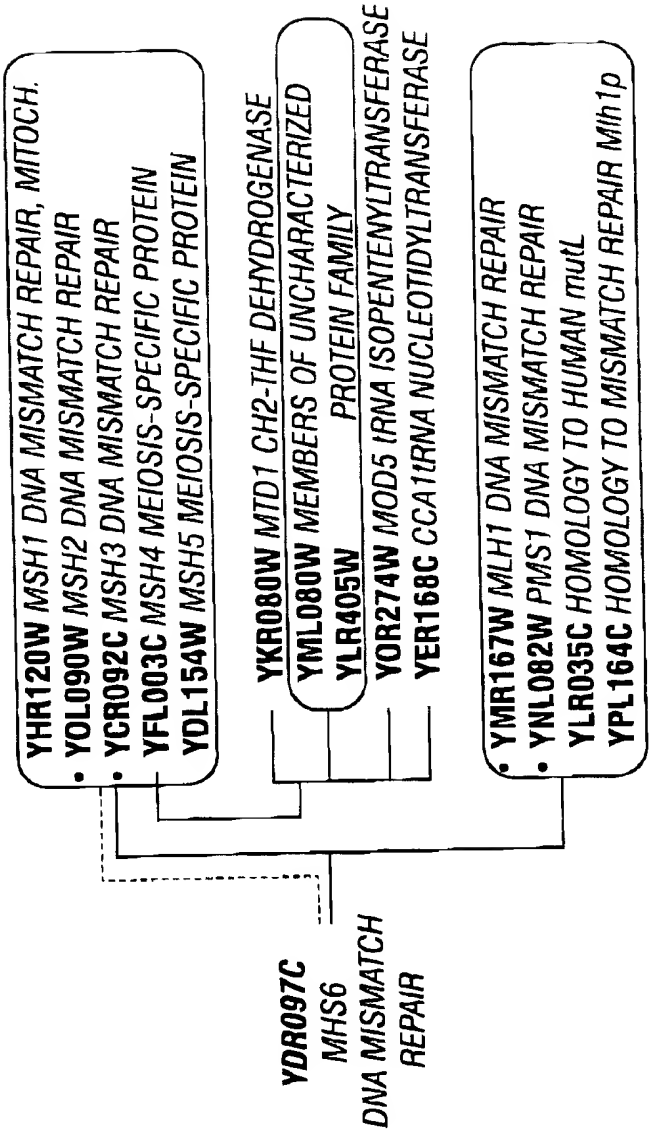
YHR120W MSH1 DNA MISMATCH REPAIR, MITOCH.
• YOL090W MSH2 DNA MISMATCH REPAIR
• YCR092C MSH3 DNA MISMATCH REPAIR
YFL003C MSH4 MEIOSIS-SPECIFIC PROTEIN
YDL154W MSH5 MEIOSIS-SPECIFIC PROTEIN
YKR080W MTD1 CH2-THF DEHYDROGENASE
YML080W MEMBERS OF UNCHARACTERIZED
YLR405W PROTEIN FAMILY
YOR274W MOD5 tRNA ISOPENTENYLTRANSFERASE
YER168C CCA1tRNA NUCLEOTIDYLTRANSFERASE
• YMR167W MLH1 DNA MISMATCH REPAIR
• YNL082W PMS1 DNA MISMATCH REPAIR
YLR035C HOMOLOGY TO HUMAN mutL
YPL164C HOMOLOGY TO MISMATCH REPAIR Mlh1p
YDR097C
MHS6
DNA MISMATCH
REPAIR

FIG. 12